

IMPROVED SPECTRAL ANALYSIS OF NEAR PERIODIC SIGNALS WITH LONG-TERM PREDICTION

Claudio BECCHETTI* and Giovanni JACOVITTI**

*Fondazione U. Bordoni, V. B. Castiglione 59, 00142 Roma, Italy e-mail: CBecc@fub.it

**Università di Rome "La Sapienza", V. Eudossiana 18, 00184 Roma, Italy

RÉSUMÉ

L'estimation du spectre de signaux quasi périodiques peut être améliorée par la sur-détermination du modèle AutoRegressive dans le contexte d'analyse à prédiction linéaire.

Dans cet étude on propose l'addition de termes de prédiction à "longue distance", ce qui se traduit en modèles d'ordre élevé avec une modeste extension des systèmes d'équations normaux. Comme dans toutes mesures de type interférométrique, cette procédure comporte la présence d'ambiguïtés qui peuvent être éliminées par l'exploitation d'information a priori. La méthode est illustrée sur deux cas simples de processus AR à bande étroite et de composants harmoniques proches.

1. INTRODUCTION

Well known parametric spectral analysis techniques for stationary time series are based on the minimisation of the linear prediction error. This criterion leads to the conventional AR solutions. These methods belong to the class of the super resolution techniques, because they fit the available data to autocorrelation model of infinite duration, so that the spectra are not affected by windowing effects.

For this reason, AR methods are well suited for high resolution estimation from relatively short sequences of data. In particular, they are employed for detection and estimation of closely spaced sinusoidal signals in noise.

Frequencies are determined by resolving a set of normal equations and then by finding the zeroes of a polynomial formed with the computed LP parameters.

It is well known that in presence of noise the resolution capability of the predictive harmonic retrieval techniques is improved by extending the number of equations L , i.e. the number of estimated autocorrelation lags beyond the minimum required by the order of the AR model (for M harmonics, the minimum value of L is $2M+1$).

From a theoretical point of view, the resolution capability of the estimate increases with the number of AR extra-poles, but it is limited in practice by the worsening of the autocorrelation estimates due to the limited set of available data and by the increased computational effort. The accuracy of such estimates can be enhanced by suppressing noise in the covariance matrix

ABSTRACT

The spectrum estimation of near periodic signals can be improved by overdetermining the AR model in a linear predictive analysis framework. In this work, "long term" prediction terms are employed. They lead to high order AR models with a small extension of the number of normal equations.

As in other interferometric methods, this technique implies ambiguity problems, which can be solved by exploiting a priori information. The technique is illustrated through two simple cases corresponding to AR narrow band processes and close harmonic pairs.

eigenspace, exploiting the a priori information about the actual number of sinusoids [1].

More recently, the problem of resolution enhancement has been approached by undersampling subband filtered versions of the input signal and by applying the TLS-Prony procedure to these single subband filtered sequences. Advantages can be achieved as single band estimation is replaced by several lower-order less computationally complex estimators [3]. In [5], the technique of undersampling the autocorrelation function has been adopted in order to modify the poles distribution in the z -plane.

In this contribution, we show that significant resolution improvement can be achieved for periodic and near periodic signals with a relatively small computational extra effort.

In essence, we increase the degree of the AR polynomials, i.e. the spectral order, incrementing by a small amount the number of normal equations L . This is done by inserting in the predictive model "long distance terms" corresponding to those lags where maxima of the autocorrelation function occur.

In other words, we explore here the possibility of inserting an interferometric concept into the AR spectra estimates.

This is inspired to the strategy employed in advanced linear prediction based speech coders like Multipulse LPC and CELP. Such coders represent the speech signals with a predictive model. The speech signal is modeled as a IIR linear system excited by a signal approximating the prediction residual. It has long been recognised that the prediction error energy is reduced by adding to the usual IIR filter modeling the reverberant effects of the acoustic cavities of the so called "vocal tract" a "long term" predictor which takes into account the basic periodicity of the "voiced sounds" [2]. The addition of long term prediction



contributions may improve performance of many other applications of linear predictive models, such as spectrum estimation, harmonic retrieval, direction of arrival (DOA) estimation of planewaves from sensor arrays, etc. . Here, we limit our analysis to spectral analysis of time series characterized by near periodic observations. More specifically, we refer to two cases of narrow band processes: high Q AR models and close sinusoid pairs.

Before entering in such applications, let us briefly illustrate the extension of linear prediction with long distance terms. We refer here to the simplest case of one long distance term.

2. EXTENDED LINEAR PREDICTION FORMULATION

The standard one-step ahead linear prediction (LP) model is described by the equation:

$$\hat{x}(n) = -\sum_{i=1}^L a_i x(n-i) \quad (1)$$

where the prediction error is defined as:

$$e(n) = x(n) + \sum_{i=1}^L a_i x(n-i) \quad (2)$$

Let us consider an Extended Linear Prediction (ELP) model, with a long term (LT) contribution:

$$\hat{x}_{LT}(n) = -\sum_{i=1}^L a_i x(n-i) - a_{LT} x(n-k) \quad (3)$$

$$e_{LT}(n) = x(n) + \sum_{i=1}^L a_i x(n-i) + a_{LT} x(n-k) \quad (4)$$

The simplest way to compute the LP model is to form a covariance matrix with standard covariance estimators:

$$\begin{bmatrix} \Phi(0,0) & \Phi(0,1) & \dots & \Phi(0,L) \\ \Phi(1,0) & \Phi(1,1) & \dots & \Phi(1,L-1) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi(L,0) & \Phi(L,1) & \dots & \Phi(L,L) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_L \end{bmatrix} = \begin{bmatrix} \sigma_e^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (5)$$

with :

$$\Phi(i, j) = \sum_{n=p-1}^{N-1} x(n-i)x(n-j), \begin{cases} 0 \leq i \leq p \\ 0 \leq j \leq p \end{cases}$$

where $x(n)$ is the input sequence of length N .

Likewise, an ELP model can be computed with the system:

$$\begin{bmatrix} \Phi(0,0) & \Phi(0,1) & \Phi(0,2) & \dots & \Phi(0,L) & \Phi(0,k) \\ \Phi(1,0) & \Phi(1,1) & \Phi(1,2) & \dots & \Phi(1,L) & \Phi(1,k) \\ \Phi(2,0) & \Phi(2,1) & \Phi(2,2) & \dots & \Phi(2,L) & \Phi(2,k) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \Phi(L,0) & \Phi(L,1) & \Phi(L,2) & \dots & \Phi(L,L) & \Phi(L,k) \\ \Phi(k,0) & \Phi(k,1) & \Phi(k,2) & \dots & \Phi(k,L) & \Phi(k,k) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_L \\ a_k \end{bmatrix} = \begin{bmatrix} \sigma_e^2 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad (6)$$

obtained from the LP model with the addition of one equation and the computation of L more autocorrelation estimates. The

resulting AR spectrum estimate of the LP params computed by (6) is:

$$P_x^{AR}(\omega) = \frac{1}{\left| 1 + \sum_{i=1}^{L-1} a_i e^{j\omega i} + a_{LT} e^{j\omega k} \right|^2} \quad (7)$$

Let us observe that in the above formulation the long term prediction term is directly inserted into an overall minimum MSE scheme. Instead, in the speech coding schemes, the long term prediction operation is cascaded with a conventional LP filter, i.e. the prediction model is formulated as follows:

$$\begin{aligned} e_{LT}(n) &= e(n) - a_{LT} e(n-k) \\ e(n) &= x(n) + \sum_{i=1}^L a_i x(n-k) \end{aligned} \quad (8)$$

In this case the overall estimate is suboptimal, but it implies the solution of separate systems of normal equations.

In both cases, we need a criterion for choosing the distance k of the long term prediction.

Following the criterion of minimising the overall prediction error energy, we should select the lag k for which:

$$E\left\{ \left| e_{LT}(n) \right|^2 \right\} = \Phi(0,0) + \sum_{i=1}^L a_i \Phi(0,i) + a_{LT} \Phi(0,k) \quad (9)$$

is minimum.

This would require in principle an exhaustive search for k . In [4] a suboptimum procedure for eliminating one equation at time from a whole set of normal equation is proposed.

Since we refer to a single long term coefficient, we adopt here the same criterion employed in the speech coding applications.

In other words, we determine k as the lag at which a relative autocorrelation maximum indicates the basic periodicity of the signals.

This means that we lean on the high predictability of the waveforms from one period to the next in order to lower the prediction error.

The AR ELP based model posses k poles. Since by hypothesis we are dealing with near periodic signals, we search for the poles closest to the unit circle. Thus, the only information we need in general, is the number of poles to be retained, which is an a priori assumption. Other criteria could be the requirement that poles lie in a given frequency interval, or a selection by means coarse (LP) estimates. In the following we adopt the highest Q criterion.

Let us illustrate such a procedure through the following examples.

3. NARROW BAND PROCESSES

Narrow band process are characterised by near periodic signals (and autocorrelation function).

The high predictability of such processes at a distance nearly corresponding to the inverse of the centroid frequency allows to formulate good predictive spectral estimates using long term prediction.

Let us consider a stationary discrete time AR(2) process. In absence of noise, a value of L equal to 3 in (5) is sufficient to accomplish a good estimate of this AR spectrum, provided that reliable acf estimates are available. This is no longer true if the AR signal is corrupted by additive noise. This is shown in fig.1 where the AR estimate in presence of white noise with different values of SNR is displayed for a 200 samples input sequence obtained by an AR(2) source with poles located in:

$$z_{1,2} = 0.995 \cdot e^{\pm j\frac{\pi}{9}}$$

In fig. 2 the result of the long term prediction using the (6) is shown. In this case, the order of the model is raised up to 18 corresponding to the peak of the estimated autocorrelation.

Elimination of the unwanted peaks is done by retaining only the pair of complex conjugate poles of the extended AR model closest to the unit circle. The result is shown in fig. 3.

Let us verify how the ELP technique applies to voice spectrum analysis. In fig. 4 the short term AR(10) spectrum of a female vowel is displayed (fundamental frequency and formants are indicated with bold lines), while the spectrum of fig. 5 is obtained with the long term prediction (with $k = 55$), and selecting the eight poles closest to the unit circle.

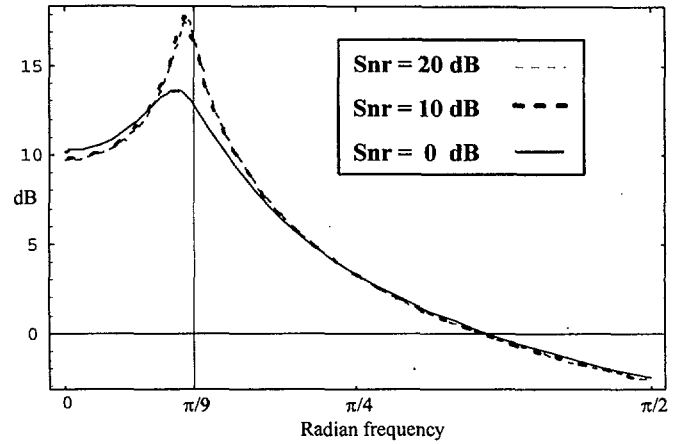


Figure 3, ELP spectral after poles cancellation

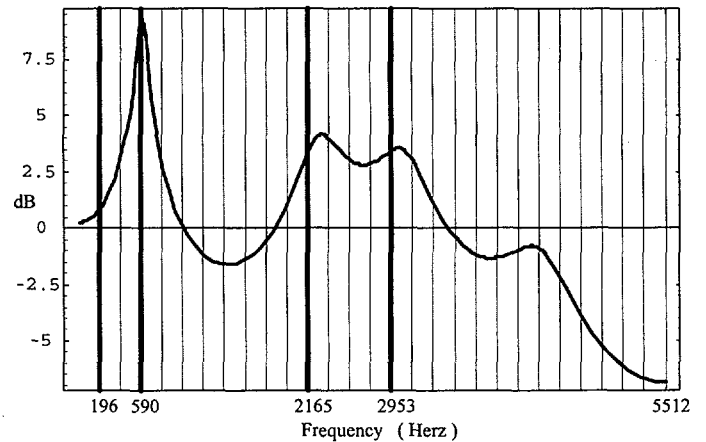


Fig. 4, AR(10) spectrum of a female "ε" vowel

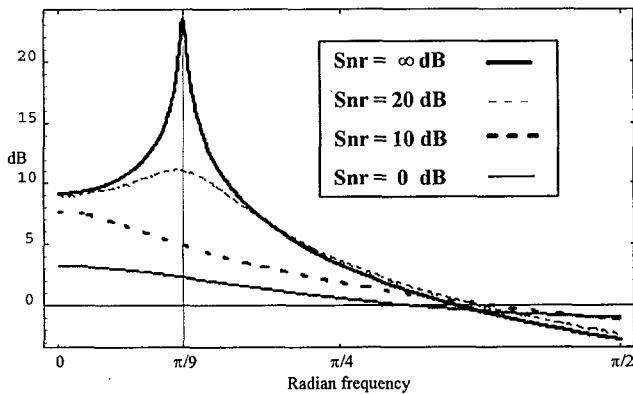


Figure 1, AR(2) spectral estimates for various SNR values

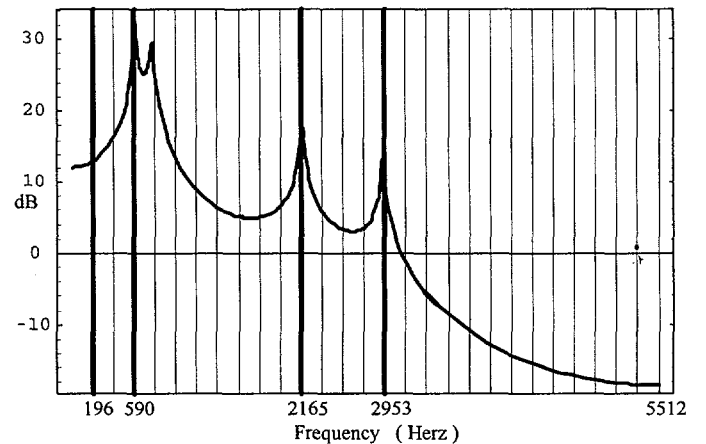


Fig. 5, ELP spectrum of a female "ε" vowel

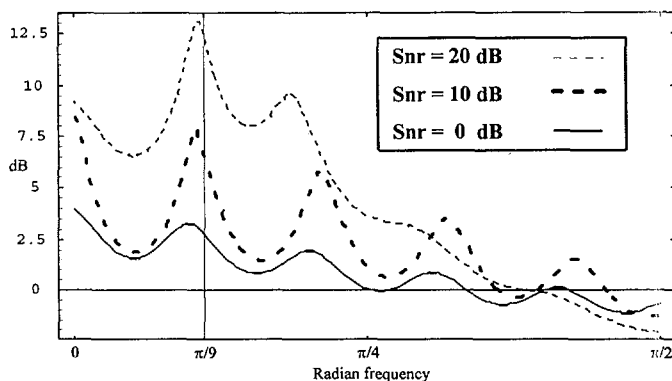


Figure 2, ELP spectral estimates for various SNR values

4. CLOSE HARMONICS

Let us now consider the following model:

$$x(n) = A_1 \cos(2\pi f_1 n + \phi_1) + A_2 \cos(2\pi f_2 n + \phi_2) + w(n) \quad (10)$$

where ϕ_1 and ϕ_2 are independent, uniformly distributed phases and $w(n)$ is an additive white noise of variance σ_w^2 . The harmonics are "close" as their frequency difference may only be discriminated by the standard AR(4) model at very high SNR values as shown in fig. 6 (we have assumed in this case that



$A_1 = A_2$ and a 200 samples input sequence with $\omega_1 = \frac{\pi}{3}$, and $\omega_2 = \frac{11}{30}\pi$.

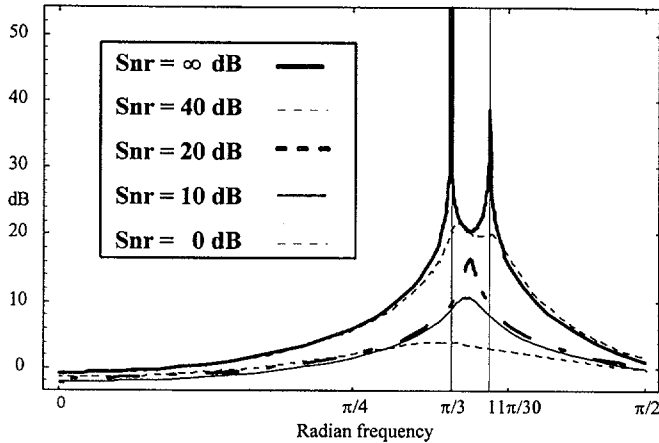


Figure 6, AR(4) spectral estimation for different SNR values

To overcome the problem of resolution loss with decreasing SNR, we may extend the model order and remove the noise eigenspace as suggested by Tuft and Kumaresan in [1] and shown in fig. 7. In this figure, a 60 order AR system projected on a four dimension space is employed.

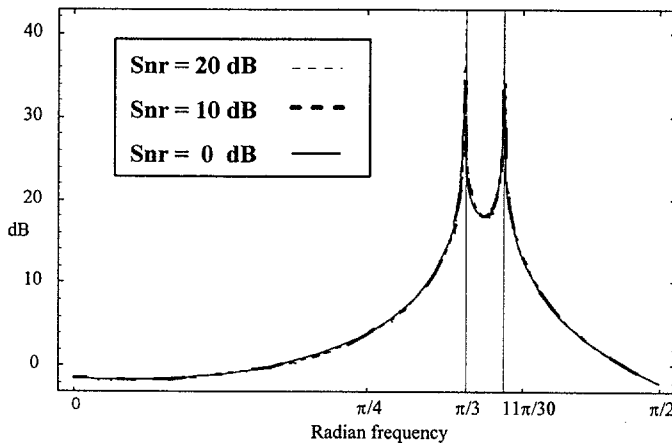


Figure 7, Tufts and Kumaresan estimation

A quite good practical result can be achieved with the much computationally cheaper ELP technique (running 63 times faster on the Mathematica platform on the average).

In fig 8, the ELP spectrum (with the correctly estimated LT value 60) is plotted showing a good resolution capability. In fig. 9 only the contributions of the poles closest to the unit circle have been displayed.

CONCLUSION

High resolution estimates of near periodic signals can be conducted with AR techniques using one more long term prediction contribution as in the "interferometric" measures. The application of this "hopped" LP order extension is that it improves the resolution of spectral estimates at moderate extra cost.

The resolution gain of this method is counterbalanced by its inherent ambiguity, which can be resolved using the a priori information (order of the system).

Further investigation is being made on the use of multiple long term contributions and on the application to eigenanalysis based spectrum analysis techniques (MUSIC).

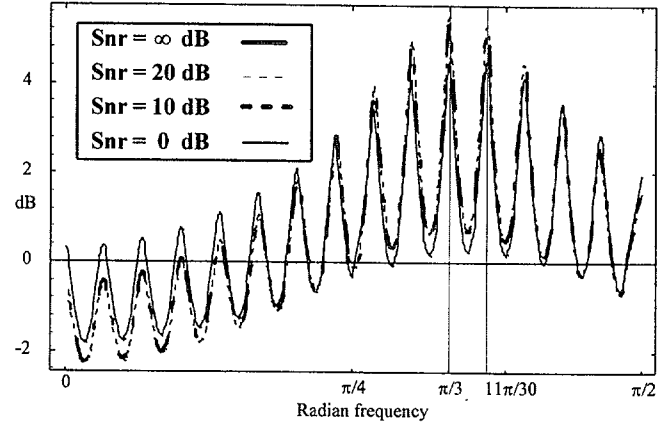


Figure 8, ELP estimation for various SNR values

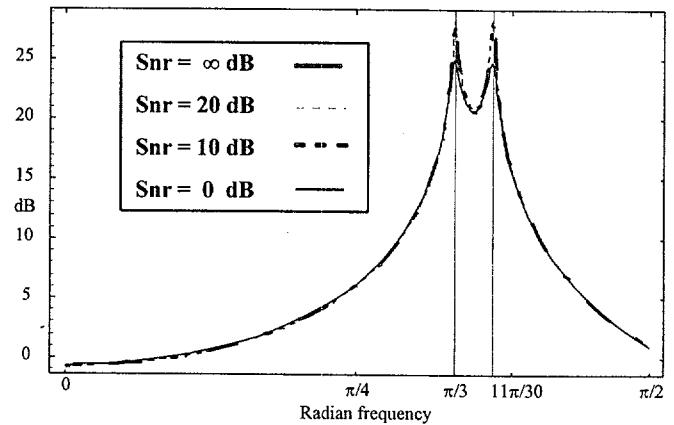


Figure 9, ELP estimation after poles cancellation procedure

REFERENCES

- [1] D. W. Tufts and R. Kumaresan: "Frequency Estimation of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood", *Proc. IEEE*, vol. 70, no. 9, pp. 975-989, Sept. 1982 .
- [2] S. Shinghal, B.S. Atal: "Amplitude Optimisation and Pitch Prediction in Multipulse Coders", *IEEE Trans. on Acoustics, Speech, and Signal Processing* vol. 37, no. 3, pp. 317-327, March 1989.
- [3] W. M. Steedly, C.J. Ying, R.L. Moses: "A Modified TLS-Prony Method Using Data Decimation", *IEEE Trans. on Signal Processing*, Vol. 42, No. 9, pp. 2292-2303, Sept. 1994 .
- [4] Rajagopalan R., et al., "Optimal Support for Linear Predictive Models", *ICIP 1994* , pp. 785-789.
- [5] M.J. Villalba, B.K. Walker, "Spectrum manipulation for improved resolution", *IEEE Transaction on Acoustic, Speech and Signal Processing*, Vol. 37, No. 6, pp. 820-831, June 1989 .